Aalborg Universitet

Interfacing Room Simulation Programs and Auralisation Systems

Møller, Henrik

Published in:
Applied Acoustics

Publication date:
1993

Document Version
Accepted author manuscript, peer reviewed version

Link to publication from Aalborg University

Citation for published version (APA):
Møller, H. (1993). Interfacing Room Simulation Programs and Auralisation Systems. *Applied Acoustics*, *36*, 333-347.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: May 05, 2023

Applied Acoustics **38** (1993) 333–347

Interfacing Room Simulation Programs and Auralisation Systems

Henrik Møller

Institute of Electronic Systems, Aalborg University, Aalborg, Denmark

(Received 9 April 1992; revised version received and accepted 6 May 1992)

ABSTRACT

Computer programs for the calculation of sound transmission in rooms are well known. Also, systems that give an aural demonstration of the calculated sound transmission have been presented. Such systems are called auralisation systems. This article describes a file format in which sound transmission programs can export data for use in auralisation systems. Utilising this data format, different sound transmission programs can be used for the same auralisation set-up, and vice versa.

1 INTRODUCTION

In recent years a number of computer programs that are capable of calculating the sound transmission in rooms have become available. The programs use a geometrical description of the room, information about the sound source(s), and knowledge about sound reflecting properties of the surfaces. The programs may be of the mirror image type,[1,2] the ray-tracing type[3–8] or combinations of these.[9–11a]

The output of a sound transmission program is normally various key figures, such as reverberation time, early decay time, early-to-late sound index (clarity), early energy fraction (Deutlichkeit), centre time and lateral energy fraction. Most programs also give reflectrograms for a given listener position.

Systems have also been presented that are able to give an audible demonstration of the sound transmission in a 'calculated' room. Various free

Applied Acoustics 0003-682X/93/$06.00 © 1993 Elsevier Science Publishers Ltd, England. Printed in Great Britain

field set-ups with loudspeakers have been used to generate sound arriving to the listener from many directions.[12-14] Also, artificially created binaural signals have been used to give the three-dimensional sound image from a simulated room.[5,6,10,15,16] The process of giving an audible demonstration of the sound transmission in a calculated room has been given the name *auralisation*.[17]

The source material for the auralisation is often a monophonic recording of a musical instrument or a human voice, made in an anechoic room. Then the system gives the auditive impression, as if the original sound source were playing in the room being simulated. By superposition it is possible to simulate several sound sources each at their position in the room.

The applications of systems including auralisation set-ups are multiple. In room acoustics they can be used to give audible demonstrations of the sound quality in concert halls and auditoriums, before the room is built. In the design of public address systems for airports, theatres, churches and other large halls, different loudspeaker units and placement strategies can be compared at the project stage. For the high-fidelity market, the quality of the sound from a loudspeaker in a variety of rooms—and in various positions in these—can be demonstrated.

The process of calculating the sound transmission in a room and the process of making this transmission audible are two separate tasks. They are linked by some information exchange, but except for that, they can be solved independently. At the *International Symposium on Computer Modelling and Prediction of Objective and Subjective Properties of Sound Fields in Rooms*, Copenhagen and Gothenburg, August 1991, the idea was put forward that the information exchange should be carried out in a standardised format. Utilising this, it would be possible to use different sound transmission programs for the same auralisation set-up, and one sound transmission program could be used with various auralisation systems. At present, a lot more sound transmission programs than auralisation programs exist, and they could gain the advantage of providing output for the relatively few auralisation systems.

Such a dividing of the tasks has already been made in a cooperation between Laboratory of Acoustics, Technical University of Denmark, and Institute of Electronic Systems, Aalborg University. The general principles of the interface between the programs were reported and agreed upon at the symposium, and the author of the present article was requested to make a more detailed description available.

Section 2 of this article contains a brief description of the principles used for auralisation together with a description of the information, which must be handed over from the sound transmission program to the auralisation system. Section 3 describes the file format in detail.

The format is very general, and in some cases programs may not be able to provide or accept data in that form. Section 4 gives two alternative descriptions that may be used, when the description of surface materials is inadequate for the standard format, or when simplified methods are used for calculation of the late part of the sound transmission.

2 AURALISATION SYSTEMS

Different systems exist for auralisation. Common to them all is that they are able to give the listener the impression of sound waves coming from all directions around him.

Two main methods are used: (i) free field methods, where the listener is exposed to sound waves really coming from the same or nearly the same directions as in the situation being simulated, and (ii) binaural methods, where only the correct sound pressures at the listener's eardrums are created, usually by means of headphones.

Both methods build up the sound as a superposition of a large number of incident sound waves, each representing a transmission path. First, the direct sound from the sound source, then the reflections from the surfaces of the room, and later on higher order reflections. The arrival time and the direction from which each sound wave arrives, are obtained from the transmission calculations, whch must also provide information about the filtering each sound wave has been exposed to during transmission (distance attenuation, air absorption, reflection attenuation, source directivity).

2.1 Free field systems with many loudspeakers

This system consists of a set-up with a large number of loudspeakers physically positioned to give the correct angles of incidence. The loudspeakers are normally put on a sphere, and the difference in arrival time is accounted for by digital delay lines. The transmission filtering is simulated by means of analogue or digital filters.

As the number of sound waves mounts to thousands, they cannot all be 'given' a loudspeaker. It is assumed that the subjective impression is undisturbed if the late reflections are given slightly incorrect angles of incidence. Therefore, only reflections up to a certain time are represented by a real loudspeaker in the correct direction. Reflections arriving after that time are sent to nearby loudspeakers. An example of a set-up of this type is described by Bech,[13] and a photograph of that set-up is shown in Fig. 1.

This type of simulation set-up has the advantage of giving very accurate simulations of the direct sound and the first reflections. Among the

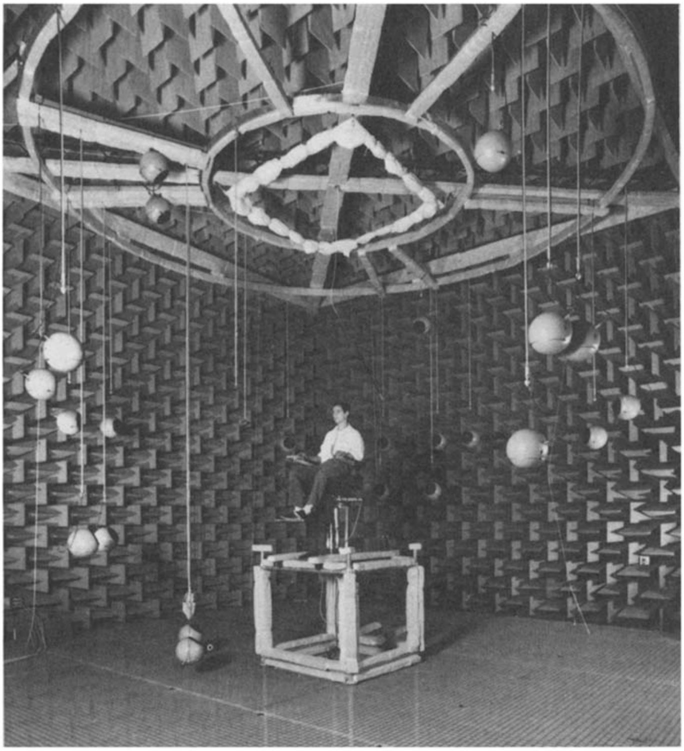

Fig. 1. Free field auralisation set-up with many loudspeakers, each in the direction of an incident wave.

disadvantages are the costs—many reproduction channels each with their loudspeaker are needed, plus a large anechoic chamber. The lack of flexibility should also be mentioned; every time a new room is to be simulated, physical changes in the loudspeaker positions are needed. Even a change in sound source or simulated listener position requires physical repositioning of loudspeakers.

2.2 Free field systems with a few loudspeakers using phantom sources

By means of two loudspeakers it is possible to give an impression of a sound coming from an imaginary sound source somewhere between the loudspeakers. The two speakers should be given the same signal, differing only in amplitude and/or in time. The imaginary sound source is called a 'phantom source', and the difference in amplitude and/or the delay between the speakers determine its position. The technique is used in traditional stereo, and it is possible to create phantom sources anywhere on the line connecting the two loudspeakers.

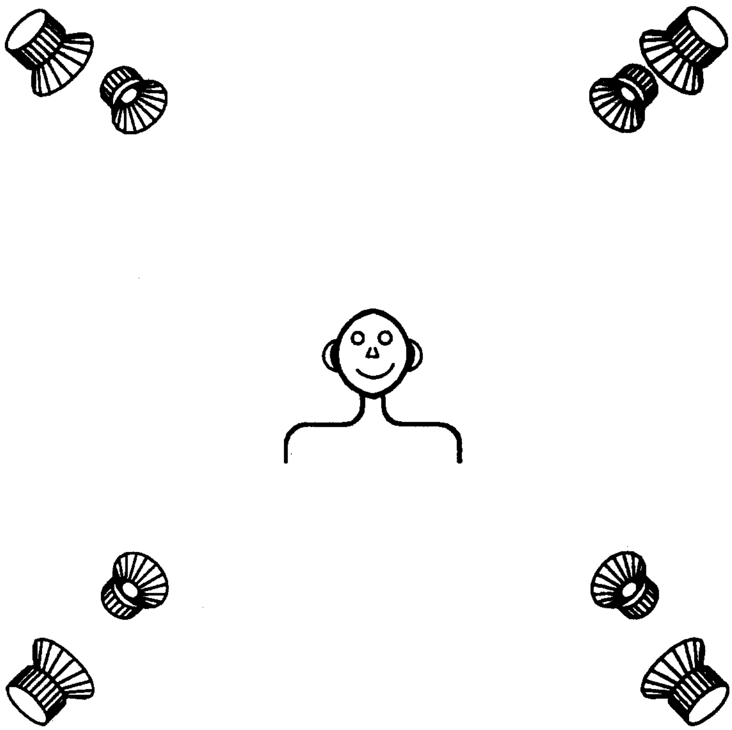

Fig. 2. Free field auralisation set-up with eight loudspeakers producing phantom sources.

Traditional stereo systems create phantom sources in the horizontal plane, but to some extent the effect works also in the vertical plane. By using three loudspeakers it is possible to create phantom sources anywhere in the space angle between them, provided they are not too far from each other. With 8–16 loudspeakers it is possible to cover the whole sphere. An example of a set-up of this type is described by Nakagawa *et al.*[14] The principle is illustrated in Fig. 2.

Costs are reduced significantly compared to the system in the previous section. A more flexible system is obtained, since the loudspeakers remain at fixed positions independent of the room being simulated. Disadvantages are inaccurate simulations of each sound wave, although it is not known at present how much this means to the total sound impression. It may also be argued that the psychoacoustic knowledge about 'three-dimensional' phantom sources is rather sparse.

2.3 Binaural systems

These systems utilise the fact that the human hearing creates the three-dimensional sound image on the basis of only two signals: the sound pressure at each of the eardrums. On its way to the two ears, a sound wave undergoes a filtering caused by diffractions around the head and body of the listener. The filtering is dependent on the direction to the sound source, and may include a delay between the two ears. The hearing is able to 'recognise' which filtering a sound wave has been exposed to, and thus determine the direction of sound incidence.

The filtering, which in the free field is carried out by the head and body, may be done electronically. For each sound wave being simulated, the sound should be sent through two filters, corresponding to the transmission to each of the two ears. This filtering can be described by the *head-related impulse responses*—$h_{left}(t)$ and $h_{right}(t)$. Head-related impulse responses are time domain representations of head-related transfer functions—and of course they are dependent on the direction of sound incidence.

If the incident sound signal is given in time as $y(t)$, then the signals $z_{left}(t)$ and $z_{right}(t)$, which should be found in the two ears can be found by convolution:

$$\begin{aligned} z_{left}(t) &= h_{left}(t) * y(t) \\ z_{right}(t) &= h_{right}(t) * y(t) \end{aligned} \tag{1}$$

A detailed description of the binaural auralisation process is given by Møller.[18]

Advantages of binaural auralisation systems are obvious. The reproduction is carried out by means of headphones, and no free field set-up is needed. Only two channels are used, and sound examples can be recorded on traditional recorders, and carried anywhere for demonstration. Among the disadvantages are that the technique requires a detailed knowledge of head-related impulse responses.

2.4 Information to be transferred from the sound transmission calculation

All three auralisation systems base their sound reproduction on knowledge of sound waves representing all relevant transmission paths from source to listener. Therefore, the main part of the data to be transferred, is a list of 'sound waves', each specified by arrival time, angle of incidence and room filtering.

The room filtering deserves some attention. It is specific for each signal path, and it should include all modifications which the signal is exposed to on its way from source to listener. This means distance attenuation, filtering due to the frequency-dependent absorption properties of surface materials, filtering due to air absorption and also filtering in accordance with the frequency response of the signal source in the transmitting direction.

The room filtering should be specified as impulse responses, given for a certain sampling frequency. Length, units, etc. will be described in the following section.

3 FILE FORMAT

The transfer of data from a sound transmission calculation program to an auralisation system is carried out by means of an ASCII data file. The file name should be given the extension ST, ST for Sound Transmission (e.g. xxxxxxxx.ST). In the case of a high sampling frequency, long reverberation time of the room and a very accurate calculation, the file size may be several megabytes, and the file can be split up into more files. If more than one file is used, the first file should have the extension ST0, and the following should be given the extensions ST1 to ST9. If more are needed, letters A to Z can be used.

The file contains four parts:

(1) an initial code, identifying the file as being in accordance with this description;
(2) a comment field at the user's disposal for describing the room, sound source, calculation method, source and receiver position, etc.;
(3) a parameter field with technical details, such as sampling frequency and others; and
(4) a list of sound waves reaching the listener.

If the file is divided into more than one file, the list of sound waves should start in the first file. The following files should start with an initial code, slightly different than in the first file, immediately followed by the continuation of the list of sound waves.

3.1 Initial code

The initial code consists of the following six ASCII characters: CUAMHX.

If more than one file is used, the number (0–9) or the letters (A–Z) from the file extension should replace the last character X. Thus, the code for the fourth file is CUAMH3.

The initial code should be followed by a line only containing the character ; (ASCII 59).

3.2 Comment field

The user is free to use this field for any identification text. The text must not contain ; (ASCII 59). There are no limits as to the length of the field, and the field may be divided into lines.

It is recommended that the lines are limited to 80 characters, and that the first two are used as text on screen display and printed output.

The comment field should be followed by a line only containing the character ; (ASCII 59).

3.3 Parameter field

The parameter field gives some technical specifications about the data. Each parameter is given a line, and it must be given by its full name in capital letters, an equality sign and the value. The parameters should be specified in the same order as shown in the following example:

SOURCE = SOUND
DESCRIPTION = COMPLETE
DOMAIN = TIME
SAMPLING FREQUENCY = 48000
NUMBER OF WAVES = 100

The SOURCE parameter accepts one of two values, SOUND or VOLTAGE. If SOUND is entered, the source is a physical sound source, and the transmission is given relative to the sound pressure in the reference direction of the sound source, at a distance of 1 m. The reference direction of a sound source is normally its frontal direction. If an anechoic recording of a sound source is used for the auralisation, it must be recorded in the reference direction (and outside the near field of the source). The dimension of the inpulse responses given for each sound wave is sound pressure/sound pressure (Pa/Pa).

If the SOURCE parameter VOLTAGE is entered, the source is one or more loudspeakers. The impulse responses will then include the frequency response of the loudspeaker(s), possibly also the power amplifier(s), equaliser(s), etc., and if more loudspeakers are used, the relative level between them. The dimension of the impulse responses is sound pressure/voltage (Pa/V).

The DESCRIPTION and DOMAIN parameters should be entered as in the example, except when the alternative descriptions given in section 4 are used.

The SAMPLING FREQUENCY parameter specifies the sampling frequency for the impulse response given for each sound wave. It should be given in Hz, and any integer can be entered. However, it is recommended that the frequencies 32 000 Hz, 44 100 Hz or 48 000 Hz are used. For a sound transmission program and an auralisation system it should be specified which sampling frequencies are implemented.

The parameter NUMBER OF WAVES gives the number of sound waves in the list. Also, this value should be of the integer type.

The parameter field should be followed by a line only containing the character ; (ASCII 59).

3.4 List of sound waves

This field contains a list of sound waves, each given by the impulse response of the transmission. The waves may or may not be ordered according to increasing arrival time at the listener position.

The description of each sound wave starts with # (ASCII 35), and on the same line follow the arrival time at the listener position relative to the time of emission, azimuth angle at the listener position, elevation angle at the listener position and number of taps in the impulse response. The arrival time should be given in seconds, and the angles in degrees according to the definitions in Fig. 3. The type of each value is real, except for the number of taps, which is integer.

The following lines should contain the individual taps in the impulse response of the sound transmission. Each tap should be given as real. There is no limit on the number of taps, and the number need not be the same for all sound waves.

For each sound wave, some comments may be inserted before the #. It is

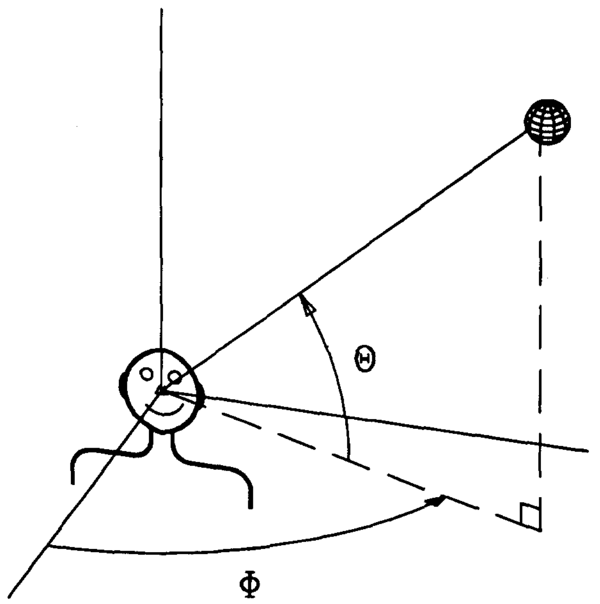

Fig. 3. Sound source and listener in a free field. Conventions for azimuth (ϕ) and elevation (θ), specifying angle of sound incidence.

a good idea (but not necessary) to end the comments with a number characterising the sound wave.

Example:

```
This is a sound wave arriving horizontally
from the left side after 50 ms
and having an impulse response with four taps
1 # 5E-2 9E1 0 4
2·31E-2
−1·28E-2
3·45E-3
−0·00 109
```

The list of sound waves should be followed by a line only containing the character ; (ASCII 59).

3.5 Example of file

An example of a (very short) file is as follows:

```
CUAMHX
;
Sound transmission in the new concert hall in XX town.
Source at podium, listener at fifth row, seat number 9.
This is a sample file showing the format of the file
transferring data from a sound transmission program
to an auralisation system.
The sound source is real (not a loudspeaker),
and four transmission paths are included.
;
SOURCE = SOUND
DESCRIPTION = COMPLETE
DOMAIN = TIME
SAMPLING FREQUENCY = 32000
NUMBER OF WAVES = 4
;
The first sound wave is the direct sound:
1 # 44E-3 −2E1 2E0 1
0·066
Then comes a reflection from the ceiling:
2 # 53E-3 −20 34 3
0·02
−0·012
```

7E-3
then from the left side wall:
3# 58·9E-3 3·5E1 0·2E1 3
0·039
−0·017
0·008
and the right side wall:
4# 0·073 52E0 3 6
−0·019
1·1E-3
2·3456E-2
−0·002
−0·012
0·003
;

4 ALTERNATIVE DESCRIPTIONS

The format described above is the general and the most accurate format. However, when following this, the demands to the sound transmission program and the auralisation system are quite high, and it is foreseen that some programs and systems will not be able to provide and accept data in this form. Therefore, two alternative formats are described. The alternatives can be used alone or in combination. For sound transmission programs and auralisation programs it should be stated whether these alternatives are supported (in some cases only these may be supported).

4.1 Frequency domain description

Specification of each transmission path in terms of an impulse response requires knowledge of the reflective properties of the surface materials, also given as impulse responses. However, the most common way of describing materials is by energy-absorption coefficients given in octaves or third-octaves. If only these are known, only the amplitude of the transmission can be found, and it is not possible to find the impulse response of the transmission path. A possible solution will be to use an approximation based on the impulse response of a minimum phase system with the correct amplitude response.

When using the described format, the conversion of the amplitude response into an impulse response should be carried out in the sound transmission program. As the materials belong to the physical sound transmission, it is quite natural that the responsibility for approximations in the description of these is given to the sound transmission program. However, some sound transmission programs may not be able to handle a time domain description, and an alternative format is therefore given.

If the frequency domain description is used, DOMAIN = TIME should be replaced by DOMAIN = FREQUENCY. SAMPLING FREQUENCY = xxxxx should be replaced by three lines, specifying the frequency resolution, the lowest frequency band and the number of bands, where data are given. The following is an example of this format:

```
SOURCE = SOUND
DESCRIPTION = COMPLETE
DOMAIN = FREQUENCY
RESOLUTION = OCTAVE
LOWEST BAND = 125
NUMBER OF BANDS = 6
NUMBER OF WAVES = 100
```

In this case, the impulse response description of each sound wave is replaced by the amplitude of the transmission given at six octave frequencies starting at 125 Hz. The resolution may also be given as THIRD-OCTAVE. The starting frequencies that may be inserted are the nominal octave or third-octave frequencies, given as integers.

In the description of each sound wave the fourth figure, normally indicating the number of taps in the impulse response, is omitted.

Example:

```
This is a sound wave arriving horizontally
from the left side after 50 ms
and having a transmission given in the frequency domain
by the amplitude at six frequencies
1 # 0·05 90 0
40E-3
32E-3
25E-3
22E-3
15E-3
8E-3
```

Drawbacks in the frequency domain description are the relatively coarse frequency resolution, and the possibly incorrect phase information.

Consequently, errors may occur in the reproduction of constructive and destructive interference (room modes are incorrect).

4.2 Incomplete description

Normally, in a sound transmission program, an accurate calculation is only carried out for the first part of the sound transmission. After some time, the number of possible transmission paths becomes very large, and most programs use an approximation. One way of doing this, is by making accurate calculations for only a fraction of the transmission paths, and adjusting the level of these to yield the correct total sound energy.

Another way of simplifying the calculation of late energy is by assuming a diffuse sound field and an exponential energy decay. When using the above-described format in combination with a such simplification, the sound transmission program should convert the diffuse sound field into discrete sound waves. Some auralisation systems use diffuse reverberation processors for the late energy, and for these it may be more practical to have the reverberation times transferred directly. An alternative format for this is given.

When using this format, the DESCRIPTION = COMPLETE line should be replaced by a DESCRIPTION = INCOMPLETE line. This should be followed by three lines specifying the frequency resolution for exponential decays, the lowest frequency band and the number of bands specified.

Example:

```
SOURCE = SOUND
DESCRIPTION = INCOMPLETE
DECAY RESOLUTION = OCTAVE
DECAY LOWEST BAND = 125
DECAY NUMBER OF BANDS = 7
DOMAIN = TIME
SAMPLING FREQUENCY = 32000
NUMBER = 100
```

THIRD-OCTAVE may also be entered as resolution, and all nominal octave and third-octave frequencies may be entered as the lowest frequency (as integers).

The data for the diffuse reverberation are entered at the end of the list of sound waves (but before the line containing ; (ASCII 59)). The format is as follows: one line is entered for each frequency; each line contains three parameters: (i) onset time of the reverberation in seconds, (ii) gain, given as the ratio between the rms levels at output and input (Pa/Pa if SOURCE = SOUND, Pa/V if SOURCE = VOLTAGE), and (iii) reverberation time.

Example:

The following describes an exponential reverberation
at six octave frequencies.
The onset time is 200 ms for all frequencies.
Note that these comments are inserted
in the same way as for the individual impulses.
0·2 0·3 1·2
0·2 0·25 1·1
0·2 0·2 1·0
0·2 0·18 0·9
0·2 0·15 0·8
0·2 0·12 0·6

The individually described sound waves and the diffuse reverberation are allowed to overlap in time.

Drawbacks in the incomplete description are the entirely diffuse sound field; there are no single waves with specific directions, and only exponential decays are reproduced. Flutter echoes, late single echoes and similar acoustic 'defects' will not be disclosed.

5 CONCLUSION

A data file interface between sound transmission programs and auralisation systems have been described. The normal form is very general and allows an accurate description. Alternative simpler formats may be suitable in some situations. It is the author's hope that this interface will be used by sound transmission programs as well as auralisation systems, and that further research and use of the systems are stimulated in this way. Comments and suggestions are readily received, together with information about systems that use or intend to use the interface. Responses received within four months from publication will be considered for a technical note, to appear in this journal.

REFERENCES

1. Kristiansen, U. R., Krokstad, A. & Follestad, T., Extending the image method to higher order reflections. *Appl. Acoust.*, **38** (1993) 195–206.
2. Staffeldt, H., Modelling of room acoustics and loudspeakers in JBL's complex array design program CADP2. *Appl. Acoust.*, **38** (1993) 179–93.
3. Lehnert, H., Strahlverfolgungsverfahren [Ray-Tracing] mit punktförmigen Quellen und Empfängern sowie idealen Strahlen. In *Forschritte der Akustik, DAGA '91*, Bochum, DPG, Bad Honnef D-5340, pp. 633–6.

4. Lehnert, H., Berechnung von langen Raumimpulsantworten durch ein schnelles Strahlverfolgungsverfahren [Ray-Tracing]. In *Fortschritte der Akustik, DAGA '91*, Bochum, DPG, Bad Honnef D-5340, pp. 637–40.
5. Heinz, R., Ein hochaufgelöstes Schallteilchenverfahren zur binauralen raumsimulation unter berücksichtigung der diffusen wandstreuung. *Acoustica*, **75** (1992) 246–55.
6. Lehnert, H., Systematic errors of the ray-tracing algorithm. *Appl. Acoust.*, **38** (1993) 207–21.
7. Maercke, D. van & Martin, J., The prediction of echograms and impulse responses within the EPIDAURE software. *Appl. Acoust.*, **38** (1993) 93–114.
8. Oguchi, K., Toyota, Y. & Nagata, M., Computer simulation technique for determining distribution of early reflections in room acoustical design. Paper presented at *International Symposium on Computer Modelling and Prediction of Objective and Subjective Properties of Sound Fields in Rooms*. August, 1991, Copenhagen and Gothenburg.
9. Vorländer, M., Simulation of the transient and steady-state sound propagation in rooms using a new combined raytracing/image-source algorithm. *J. Acoust. Soc. Amer.*, **86** (1989) 172–8.
10. Dalenbäck, B.-I., Auralization based on image source modelling augmented by ray tracing and diffuse reflections. Paper presented at the *International Symposium on Computer Modelling and Prediction of Objective and Subjective Properties of Sound Fields in Rooms*. August, 1991, Copenhagen and Gothenburg.
11. Naylor, G. M., ODEON—Another hybrid room acoustical model. *Appl. Acoust.*, **38** (1993) 131–43.

11*a*. Heinz, R., Binaural room simulation based on an image source model with addition of statistical methods to include the diffuse sound scattering of walls and to predict the reverberant tail. *Appl. Acoust.*, **38**(2–4) (1993) 145–59.

12. Kleiner, M., Studies of simulation techniques in room acoustics. Report 7809, Department of Building Acoustics, Technical University of Gothenburg, Sweden.
13. Bech, S., Electroacoustic simulation of listening room acoustics; psychoacoustic design criteria. Preprint 2989, 89th convention of Audio Engineering Society, Los Angeles, CA, USA.
14. Nakagawa, K., Miyajima, T. & Tahara, Y., An improved geometrical sound field analysis in rooms using scattered sound and an audible room acoustic simulator. *Appl. Acoust.*, **38** (1993) 115–29.
15. Svensson, P., Auralization of reverberation enhancement systems. Paper presented at the *International Symposium on Computer Modelling and Prediction of Objective and Subjective Properties of Sound Fields in Rooms*. August, 1991, Copenhagen and Gothenburg.
16. Xiang, N. & Blauert, J., Binaural scale modelling for auralization and prediction of acoustics in auditoria. *Appl. Acoust.*, **38** (1993) 267–90.
17. Kleiner, M., Svensson, P. & Dalenbäck, B.-I., Auralization: Experiments in Acoustical CAD. Preprint 2990, 89th Convention of the Audio Engineering Society, Los Angeles, CA, USA.
18. Møller, H., Fundamentals of binaural technology. *Appl. Acoust.*, **36** (1992) 171–218.